Design of a Biosignal Based Stress Detection System Using Machine Learning Techniques

[1]U. SRINIVASULU, [2]K.SAI, [3]CH. SAI LIKITHA, [4]T. PREM, [5]P. SADIQ ALI.

[1]Professor (Assistant) of ECE Department, SIETK, PUTTUR, AP
[2, 3, 4, 5] UG SCHOLAR OF ECE Department, SIETK, PUTTUR, AP

[1, 2, 3, 4, 5]Department of Electronics & Communication Engineering
Siddharth Institute of Engineering & Technology (Autonomous),
Puttur-517583, Andhra Pradesh

utchuru.srinivas@gmail.com, kattasai0504@gmail.com,
sailikitha982@gmail.com, thalariprem126000@gmail.com
sadiqalipatnool@gmail.com

***Abstract*— This study represents a design of a detection system of stress through machine learning using some available bio signals in human body. Stress can be commonly defined as the disturbance in psychological equilibrium. Stress detectionis one of the major research areas in biomedical engineering as proper detection of stress can conveniently prevent many psychological and physiological problems like cardiac rhythm abnormalities or arrhythmia. There are several bio-signals available (i.e. ECG, EMG, Respiration, GSR etc.) which are helpful in detecting stress levels as these signals shows characteristic changes with stress induction. In this paper, ECG was selected as the primary candidate because of the easily available recording (i.e. several mobile clinical grade recorders are available now in the market) and ECG feature extraction techniques. Another advantage of ECG is that respiratory signal information can also be detected form ECG which is known as EDR (ECG derived Respiration) without having separate sensor system for respiration measurement. Features of ECG signals are distinctive and collection of the signals is cost-efficient. From ECG we derived RR interval, QT interval, and EDR features for the development of the model. For the implementation of a supervised machine learning (SVM) method in MATLAB, Physionet's "drivedb" database was used as the training dataset and validation. SVM was chosen for classification, as there are two classes of labeled data; 'stressed' or 'non-stressed'. Several SVM model types were verified by changing the feature number and Kernel type. Our results showed an accuracy level of 98.6% with Gaussian Kernel function and using all available features (RR, QT and EDR), which also emphasizes the importance of respiratory information in stress detection**

through Machine Learning.

Keywords— stress detection, arrhythmia, ECG Derived Respiration (EDR), Machine Learning, MATLAB

I. INTRODUCTION

Stress can be commonly interpreted as the disruption in normal psychological equilibrium. When a person is unable to balance between the demands that are placed on him/her and his/her ability to cope with them, then it causes pressureon mental health which creates stress. There are two kinds of stresses. Eustress is defined as positive stress [1] and distress is when stress has negative impact in life. Mental professionally. Stress can lead to various kinds of health problems especially several forms of cardiovascular diseases [2-4]. Figure 1 describes graphically how different form of stress affects two branches of Autonomic Nervous system and their after effects such as arrhythmogenesis [7]. The sympathetic nervous system (SNS) is stimulated by physiological stress and this can result in arrhythmias especially Ventricular tachyarrhythmias [3-7]. Many physiological parameters are also affected by the activation of SNS such as pupil dilation, skin conductivity, respiration and muscle activity [10-13]. In addition, different parameters of an ECG such as RR, ST, QT intervals as well as heart rate variability (HRV) are affected by stress [14-18]. Reduced heart rate variability, increased QT dispersion, and changed respiratory pattern are found to be led by stress [19-20]. People with these changes have the utmost risk for emerging fatal ventricular arrhythmias [6-7]. Therefore, detection or identification of stress is crucial in preventing the initiation of cardiac complexities.

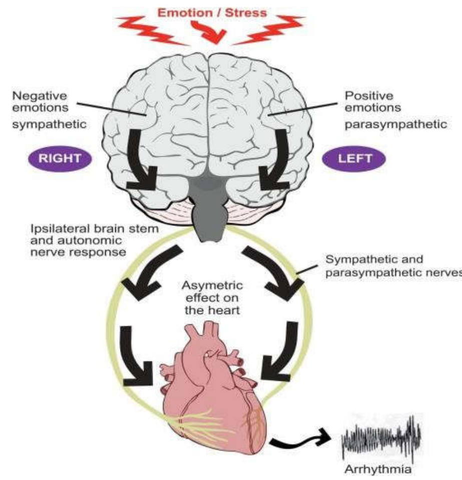


Figure 1: How Stress induce Arrhythmia [7]

To identify stress, questionnaire-based counseling is a conventional method that is commonly used. However, this method may seem time-consuming as well as costly as a psychological expert is needed to assess whether a person is stressed or not. That is Bio-signal based detection techniques are getting popularity that can determine real- time stress and can effectively save time and cost. Many researchers have studied different bio signals to identify real-time stress such as ECG, galvanic skin response (GSR) [8], electromyogram (EMG) [9-11], skin temperature (ST) [12], skin conductivity [13], respiratory rate and respiration amplitude [14], blood pressure etc. ECG signal is widely used in stress-based researches, as collection of ECG data is simpler, distinctive and affordable now due to advances in effective mobile ECG recorders. In addition, a lot of information besides heart rate variability (HRV) parameters including respiratory pattern can be extracted from ECG, which makes it an ideal choice for stress detector [15- 16,21].

Until now, almost all other methods have only used HRV (i.e. RR interval based different statistical features) as the main features to detect stress [10-16]. In our study, we used three crucial features form ECG: QT interval, RR interval and EDR (ECG Derived Respiration) to detect stress. People having stressed condition may be identified by the prolongation of QT interval. The sympathetic nervous system is affected by stress [17]. QT intervals of ECG during stressed condition were investigated and the outcome showed that QT intervals are shortened when a person is stressed [18]. Respiration and the cardiovascular system are strongly interconnected with each other [19]. Stressed individual has a faster and shallower breathing pattern [20]. Respiratory rate and respiration amplitude are associated with SNS activity, which directs to stress [19]. EDR is used as a substitute for the original respiratory signal as they show similar behavior [21]. Use of EDR removes the necessity of having extra respiration measurement system. Supervised machine learning method was chosen to detect stress by using Physionet's "drivedb" database as training data [22]. To the best of our knowledge, little work has been carried out on stress detection through machine learning using bio-signals and previous works have not been comprehensively considered due to low model performance [13]. The following sections describe the model formation and the model performance characteristics analysis.

2.DATA FOR MODELLING

For training the model, Stress Recognition in Automobile Drivers database ("drivedb") available at Physionet (www. Physionet.org) was used where some healthy subjects' normal and stressed condition's ECG data were recorded [22]. In this database, an experimental protocol was designed and verified for the detection of stress due to driving in heavy traffic condition form physiological signals of the healthy subjects. According to the designed procedure, subjects were driving car following a set route and their physiologic reactions were monitored by analyzing several recorded physiological signals like Electrocardiogram (ECG), Electromyogram (EMG), skin conductivity and respiration [22]. The driving protocol was designed to take the driver in different road conditions with variable traffic such that different levels of stress were likely to occur. A total number of 15 healthy subject's data were taken. In our study, we have used 5 minutes of ECG signal during resting and high-stress condition. The resting conditions of the drivers were considered as "Not Stressed" condition and driving in heavy traffic conditions were considered as "Stressed" condition.

The ECG signals were filtered for baseline removal and features were extracted using the same methodology as described in details in [21]. In this algorithm, to detect different peaks of ECG (i.e. R wave peak, Q wave pint, T wave peak and T wave end) RR interval and QT interval time series were formed for the whole length of ECG. EDR signal was derived as the variation of R wave amplitude. Almost 2500 observations for both resting and stressed condition (i.e. RR, QT and EDR data points) were used to train the model.

3.METHODOLOGY AND MODELING

The modeling and validation were done according to the block diagram in figure 2. Classification Learner app from MATLAB's Machine Statistics and Machine Learning Toolbox was used to train and validate the model.

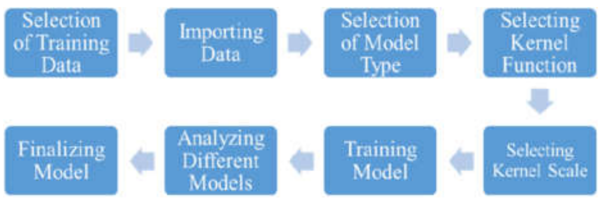


Figure 2: Block Diagram of SVN Model analysis

In this study, Supervised Machine Learning has been used todetect stress. ECG features that were chosen to detect stress are QT interval, RR interval and ECG Derived Respiration. By analyzing them, the model will decide if the subject is stressed or relaxed. As there are exactly two class labels in our data, (i.e. stressed and Not stressed or relaxed) we have chosen Support Vector Machine (SVM) for classification and detection of stress.

Qtend	RRInterval	EDR	Decission
0.47782	0.7399194	0.75308	Not Stressed
0.49798	0.8165323	0.76355	Not Stressed
0.41331	0.8790323	0.81982	Not Stressed
0.4254	0.8266129	0.8214	Not Stressed
0.41331	0.7721774	0.74514	Not Stressed
0.40927	0.8608871	0.79797	Not Stressed
0.42339	0.9052419	0.81863	Not Stressed
0.41331	0.7842742	0.79239	Not Stressed
0.48589	0.7600806	0.74238	Not Stressed
0.41331	0.8125	0.79313	Not Stressed
0.41331	0.8326613	0.81291	Not Stressed
0.37702	0.7762097	0.76596	Not Stressed
0.39315	0.7741935	0.65084	Not Stressed
0.40927	0.9254032	0.79558	Not Stressed
0.37097	0.8870968	0.79862	Not Stressed
0.49194	0.7903226	0.76279	Not Stressed
0.4254	0.7741935	0.75647	Not Stressed

(a)

Qtend	RRInterval	EDR	Decission
0.3629	0.9012097	1.48158	Stressed
0.36492	0.9193548	1.54413	Stressed
0.36694	0.8245968	1.45752	Stressed
0.35484	0.8830645	1.47348	Stressed
0.36492	0.9637097	1.50651	Stressed
0.37097	0.8447581	1.37285	Stressed
0.3629	0.8870968	1.48037	Stressed
0.36492	0.8991935	1.49324	Stressed
0.3629	0.8165323	1.47152	Stressed
0.36492	0.9012097	1.50668	Stressed
0.35887	0.9495968	1.54069	Stressed
0.34476	0.8447581	1.47174	Stressed
0.35685	0.9112903	1.48052	Stressed
0.36694	0.9798387	1.53021	Stressed
0.36895	0.9012097	1.51735	Stressed
0.36694	0.8830645	1.45438	Stressed
0.36492	0.9334677	1.50834	Stressed

(b)

Figure 3: Training data selection

Support Vector Machine (SVM) can be used for both discrete and continuous data sets. Nevertheless, it is mostly used for discrete data sets, i.e. Classification Technique. In this method, the data is plotted into an n-dimensional plane where n is the number of features. Then classification is performed by finding the hyperplane that differentiates the two classes. Figure 3 shows the Data and decision variables processed for Model Training.

First, the model was trained with three SVM typess (Linear, Quadratic and Cubic) using the default kernel functions and the cross validation scheme was chosen as the holdout validation with a degree of 50% in the Classification Learner App. All of the three features i.e. QT interval, RR interval, and ECG Derived Respiration (EDR) were used. By analyzing the scatter plots, confusion matrix, and ROC curves the accurate model can be chosen. Model accuracy depending on different types of SVM is represented in Table

1. Further modifications can be done with the model to find accurate results.

Table 1: Model accuracy for different types of SVM model using all the features and the default kernel function in MATLAB

Model Name	**Features Used**	**Accuracy**
Linear SVM		52.6%
Quadratic SVM	QT interval, RR interval, EDR	88.6%
Cubic SVM		97.2%

To analyze the effect of different features, the model was then trained using only one feature to understand if it is possible to detect stress accurately. A significant drop in the model accuracy was found if we use only one feature as shown in Table 2. If we train the model with only two features by unchecking any of the three features from "Feature Selection" option of Classification Learner and train the model again, the model accuracy becomes low compared to the previous one. Model accuracy differences by training with only two features can be noticed from Table 3.

Table 2: Model accuracy by using only one feature

Model Name	QT Interval	RR Interval	EDR
Linear SVM	50.5%	61.3%	48.6 %
Quadratic SVM	30.6%	40.8%	48.7 %
Cubic SVM	61.5%	54.5%	49.0 %

Combination of QT interval and RR interval was found To give an accuracy of almost 95%. Other models failed to generate comparable amount of accuracy using combinationof two features (Table 3). Almost all of the model-basedstress detection techniques until now have used only onefeature, mostly RR interval to detect stress as found in theliterature. But their performance is not acceptable forbiomedical applications (i.e. Accuracy level less than 95%).Therefore, in this study we trained the model with QTinterval, RR interval and ECG Derived Respiration (EDR)separately and analyzed the results. We have chosen EDR(ECG Derived Respiration) as an alternate for respirationsignal as they show similar properties and be used assuccessfully in many previous studies [21]. No externalsensors were used to record respiration signal. To the best ofour knowledge, no previous study has used respiratoryinformation from ECG (ECG Derived Respiration or EDR)and QT interval in stress detection using machine learningtechniques.

Table 3: Model accuracy by using combination of two features used for modeling

Model Name	QT, RR	QT, EDR	RR, EDR
Linear SVM	50.6%	52.6%	61.5%
Quadratic SVM	83.5%	86.3%	51.2%
Cubic SVM	94.9%	89.2%	53.1%

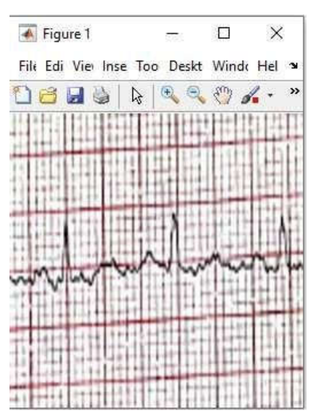
"Kernel Function" can also be tuned to improve the model performance. Therefore, we have trained the models with different kernel functions. Linear and Quadratic kernel functions did not show good results. Only Gaussian and Cubic kernel function types showed promising results in detecting stress using all three features of ECG as shown in Table 4.

Table 4: Model performance by changing Kernel Functions

Model Type	Kernel Function	Accuracy
Linear SVM	Gaussian	98.6%
	Cubic	97.2%
Quadratic SVM	Gaussian	98.6%
	Cubic	97.1%
Cubic SVM	Gaussian	98.6%
	Cubic	97.2%

RESULT

Fig: Input Image



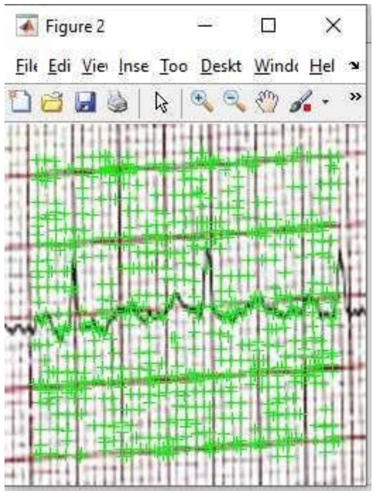


Fig: Features Image

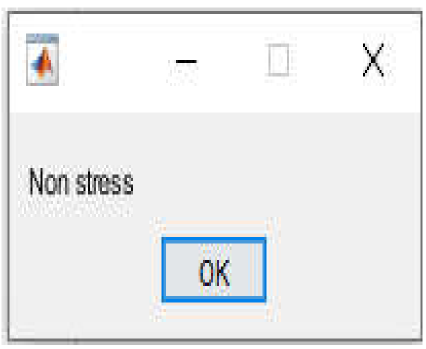


Fig: Predicted output

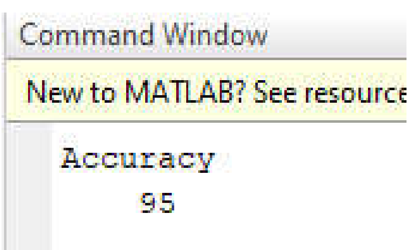


Fig: Accuracy

DISCUSSION AND CONCLUSION

Table 1 shows the model accuracies using three features of ECG. From Table 1, it was found that the model cannot be used for detection of stress if we train it with linear SVM, because it shows only 52.6% accuracy. We analyzed the confusion matrix of linear SVM and noticed that, the number of correctly predicted stressed and relaxed data is only 44% and 61% respectively. On the other hand, other SVM models i.e. Quadratic SVM, and Cubic SVM showed higher accuracy. Among them, Cubic SVM shows the highest accuracy of 97.2% with default Kernel function (i.e Cubic Kernel type) the number of correctly predicted stressed and relaxed data of 98% and 97% respectively. Therefore, we can say that, without changing the default kernel function is in Classification Learner APP in MATLAB, Cubic SVM model is the best to detect stress.

If QT interval, RR interval and ECG Derived Respiration (EDR) are used separately to detect stress, the model shows very poor accuracy as shown in Table 2. Therefore, it can be concluded that the model performance will not be acceptable if we use only one feature of the ECG signal for stress detection.

If two features were used to train the model, the model accuracies were also found to be very low except of Cubic SVM for QT and RR interval. Changes with model accuracy after deducting one feature from the list can be observed in Table 3. In all cases, Linear SVM is showing very poor accuracy in detecting stress. Even if Cubic SVM is showing a bit of higher accuracy with QT interval and RR interval- based models, other feature combinations produce unacceptable performance. RR interval and EDR based

.

model was showing significantly low accuracy than other feature-based models. Therefore, it can be concluded that the model will be more accurate if we use three features of ECG signal rather than two. Moreover, as long as QT interval and Respiratory information (EDR) are used as one of the two features, the model performs better than the models without QT interval. Therefore, it can be said that, QT interval and EDR are important parameters to detect stress as established by some recent studies [22].

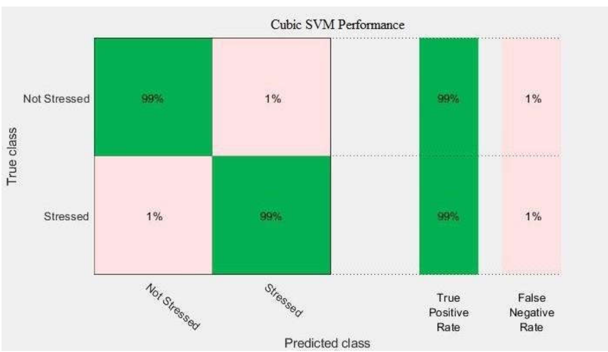


Fig 4: Confusion Matrix for the Cubic SVM using Gaussian Kernel function

We have also tried training different models by changing their kernel functions. During these training, three features of the ECG signal were used. We can easily compare the values of Linear SVM models from 1 with Table 4 and say that the model becomes very accurate if we use different kernel functions other than the default one (i.e. Linear Kernel) to modify the model. Moreover, it was found that all the models with level of high accuracy as shown in Table 4. In this case, he best model was chosen by analyzing their confusion matrix showing true positive and false negative rates. Cubic SVM with kernel function Gaussian having an accuracy of 98.6% was chosen as the best model to detect stress. It was chosen because it has correctly predicted stressed and relaxed data of 99% both where all other models with Gaussian kernel function have correctly predicted stressed and relaxed data of 99% and 98% respectively (Figure 4 and 5).

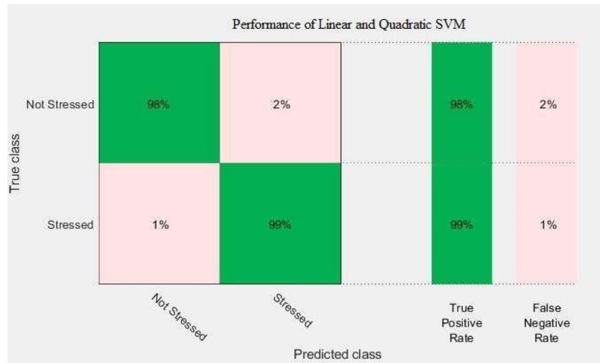


Fig 5: Confusion Matrix for the Linear and Quadratic SVM Gaussian kernel function

In this paper, different models to detect stress has been trained using multiple ECG features such as QT interval, RRinterval, and EDR. This method to detect stress from ECG signal can help an individual to assess one’s psychological condition as well as physical condition, from which he/she will be able to take necessary precautions. It was also concluded that, the more features we use, the more accurate the model becomes.

REFERENCES

[1] Fevre, Mark Le; Kolt, Gregory S.; Matheny, Jonathan,."Eustress, distress and their interpretation in primary andsecondary occupational stress management interventions:which way first?". *Journal of Managerial Psychology*,2006, 21 (6): 547 565. doi:10.1108/02683940610684391.

[2] Das S, O'Keefe J. “Behavioral cardiology: recognizingand addressing the profound impact of

psychosocial stresson cardiovascular health". Curr Atheroscler Rep. 2006;8:111-8.

[3] Shusterman V, Aysin B, Gottipaty V, Weiss R, Brode S,Schwartzman D, Anderson KP. "Autonomic NervousSystem Activity and the Spontaneous Initiation ofVentricular Tachycardia." ESVEM Investigators.Electrophysiologic Study Versus ElectrocardiographicMonitoring Trial. J Am Coll Cardiol. 1998; 32:1891-9.

[4] Brodsky MA, Sato DA, Iseri LT, Wolff LJ, Allen BJ."Ventricular Tachyarrhythmia Associated with Psychological Stress." JAMA. 1987; 257:2064-7.

[5] Leenhardt A, Lucet V, Denjoy I, Grau F, Ngoc DD,P. "Catecholaminergic polymorphic ventricula tachycardia in children. A 7-year follow-up of 21 patients."Circulation. 1995;91:1512-9.

[6] Brunckhorst CB, Holzmeister J, Scharf C, Binggeli C,Duru F. "Stress, depression and cardiac arrhythmias." TherUmsch. 2003;60:673-81.

[7] Taggart, Peter & Boyett, Mark & Logantha, Sunil Jit &Lambiase, Pier. (2011). "Anger, Emotion, and Arrhythmias:From Brain to Heart. Frontiers in physiology." 2. 67.10.3389/fphys.2011.00067.

[8] U. Lundberg, R. Kadefors, B. Melin et al., "Psychophysiological stress and emg activity of the trapezius muscle," International Journal ofBehavioral Medicine, vol. 1, no. 4, pp. 354–370, 1994.

[9] J. Wijsman, B. Grundlehner, J. Penders, and H. Hermens, "Trapezius muscle EMG as predictor of mentalstress," in Proceedings of the 1st Wireless HealthConference (WH '10), pp. 155–163, ACM, October 2010.

[10] Crescentini C., Chittaro L., Capurso V., Sioni R., Fabbro F., "Psychological and physiological responses tostressful situations in immersive virtual reality: Differencesbetween users who practice mindfulness meditation andcontrols", Computers in Human Behavior, Vol. 59, June2016, pp. 304-316

[11] J. Taelman, S. Vandeput, A. Spaepen, and S. V. Hu_el,"Influence of mental stress on heart rate and heart ratevariability," in Proceedings 4th European Conference of theInternational Federation for Medical and BiologicalEngineering (IFMBE '09), pp. 1366–1369,2009.

[12] T.S. Lorig, "The Respiratory System," *Handbook of Psychophysiology*, J.T. Cacioppo, L.G. Tassinary, and G.G.Berntson, eds., Cambridge Univ. Press, 2007

[13] NermineMunla, Mohamad Khalil, Ahmad Shahin, "Driver Stress Level Detection Using HRV 2019 International Conference on Robotics,Electrical and Signal Processing Techniques (ICREST) Analysis"International Conference on Advances in Biomedical Engineering (ICABME), 2015.

[14] Ya Jun Yu, Zhan Yang, Beom-Seok Oh, Yong KiangYeo, Qinglai Liu, Guang-Bin Huang and Zhiping Lin,"Investigation on Driver Stress Utilizing ECG signals with On-board Navigation Systems in Use", 14th International Conference on Control, Automation, Robotics & Vision, Phuket, Thailand, 13-15th November 2016 (ICARCV), 2016.

[15] A. Bhide, R. Durgaprasad, L. Kasala*et al.*, "Electrocardiographic changes during acute mental stress"*International Journal of Medical Science and Public Health(Online First)*, vol. 5, no. 5, 2016.

[16] M. Huerta-Franco, F. Vargas-Luna, and I. Delgadillo-Holtfort, "Effects of psychological stress test on the cardiacresponse of public safety workers:

alternative parameters toautonomic balance," in *Journal of Physics: ConferenceSeries*, vol. 582, no. 1, 2015, p. 012040

[17] G. Andr´assy, "The effect of different stressors on theQT interval and the T wave," Orvosi Hetilap, vol. 150, no.10, pp. 447–457, 2009.

[18] Paavonen KJ, Swna H, Pippo K, Hokkanen L et al,"Response of the QT interval to mental and physical stressin types LQT1 and LQT2 of the long QT syndrome", Heart.2001 Jul; 86(1):39-44

[19] T.S. Lorig, "The Respiratory System," Handbook ofPsychophysiology, J.T. Cacioppo, L.G. Tassinary, and G.G.Berntson, eds., Cambridge Univ. Press, 2007

[20] S.D. Kreibig, "Autonomic Nervous System Activity inEmotion: A Review," Biological Psychology, vol. 84, no. 3,2010, pp. 394–421.

[21] Imam, M. H., Karmakar, C. K., Khandoker, A. H., &Palaniswami, M., "Effect of ECG-derived respiration (EDR)on modeling ventricular repolarization dynamics in different physiological and psychological conditions". *Medical &Biological Engineering & Computing, 52*(10),2014, 851-860. doi: 10.1007/s11517-014-1188-0.

[22] J. A. Healey and R. W. Picard, "Detecting stress duringreal-world driving tasks using physiological sensors," IEEE Transactions on Intelligent Transportation Systems, vol. 6,pp. 156-166, Jun 2005.